Multivariate Kurtosis for Measuring Image Sharpness

Nien Fan Zhang[1], Michael T. Postek[1], Robert D. Larrabee[1] and Andras E. Vladar[1]

[1] National Institute of Standards & Technology, 100 Bureau Dr., Stop 8980, MD 20899-8980, USA

Abstract: In industry applications, such as automated on-line semiconductor production, users of scanning electron microscope (SEM) metrology instruments would like to have these instruments function without human intervention for long periods of time, and to have some simple criterion (or indication) of when they need servicing or other attention. At the present time, no self testing is incorporated into these instruments to verify that the instrument is performing at a satisfactory performance level. Therefore, there is a growing realization of the need for the development of a procedure for periodic performance testing. Postek and Vladar (1996) published a procedure which was based on the objective characterization of the spatial Fourier transform of the SEM image of a test object for this purpose and the development of appropriate analytical algorithms for characterizing sharpness. In this paper, an alternative approach based on the multivariate kurtosis is proposed to measure the sharpness of SEM image.

Scanning electron microscopes are being utilized extensively in the production environment. Since these instruments are approaching full automation, objective diagnostic procedures must be implemented to ensure data and measurement fidelity. One approach to this issue is the sharpness technique. It is known that the low-frequency changes in the video signal contain information about the large features and that the high-frequency changes carry information of finer details. When an SEM image has fine details at a given magnification, namely, when there are more high-frequency changes in it, we say it is sharper.

Since an SEM image is composed of a two-dimensional array of data, it can be expressed as $I(u,v)$, ($u = 1, \ldots, n$ and $v = 1, 2, \ldots, n$), where u and v are spatial indices. The corresponding two-dimensional finite Fourier transform is

$$f(x,y) = \sum_{u=1}^{n} \sum_{v=1}^{n} I(u,v) e^{i(xu+yv)} \tag{1}$$

where the spatial frequencies x and y have indices from $-N$ to N. Based on $f(x,y)$, we observe that when an SEM image is visually sharper than a second image, the high spatial frequency components of the first image are larger than those of the second. The following is an illustrative example. Part a of the attached figure shows the performance of a SEM on a heavy gold-coated oxide test sample at low accelerating voltage. This micrograph was taken following a tip change. This image appears to be far less sharp and lacking in resolution when it compared with a similar micrograph (Part c of the figure) taken when the same instrument was operating more optimally. Parts b and d show the magnitude distribution of the two-dimensional spatial Fourier transform for the images in Parts a and c, respectively. From these figures it is clear that for the sharper image in Part c of the figure, its cone in Part d, which represents the magnitude distribution of the Fourier transform of the image, is wider than that in Part b of the image in Part a of the figure.

For a given univariate random variable Z with mean μ_z and finite moments up to the fourth, the kurtosis is defined as

$$\beta_2 = \frac{m_4}{m_2^2}$$

where m_4 and m_2 are the fourth and second central moments respectively, that is

$$m_4 = E[(Z - \mu_z)^4] \text{ and } m_2 = \sigma_z^2 = E[(Z - \mu_z)^2].$$

For any univariate normal distribution, $\beta_2 = 3$. Therefore the value of β_2 can be compared with 3 to determine whether the distribution is "peaked" or "flated-topped" relative to a Gaussian. Note that kurtosis is a dimensionless ratio.

Four separate distribution density functions with zero mean and unit variance were compared by Kaplansky (1945) to illustrate the properties of kurtosis. His results show that the smaller the kurtosis, the flatter the top of the distribution. Finucan (1964) also discussed the interpretations of kurtosis. The conclusion is that the distribution with smaller kurtosis is more flat-topped or has a large shoulder than that with larger kurtosis.

Note that since kurtosis is a dimensionless ratio of the moments, a value of kurtosis can be calculated for any positive function when the area underneath the function curve is finite and when the curve has finite moments up to the fourth. Let $w = f(z)$ be such a discrete univariate function. Then

$$\beta_2 = \frac{\sum_{j=1}^{n} (z_j - \mu_z)^4 f(z_j)}{[\sum_{j=1}^{n} (z_j - \mu_z)^2 f(z_j)]^2}$$

where $\mu_z = \sum_{j=1}^{n} z_j f(z_j)$.

Multivariate kurtosis has been proposed by Mardia (1970). Let W be a p-dimensional random vector with finite moments up to the fourth. Let μ be the mean vector and Γ be the covariance matrix of W. The kurtosis of W is defined by

$$\beta_{2,p} = E[(W - \mu)^T \Gamma^{-1} (W - \mu)]^2,$$

where T denotes the transpose of a vector. When $p = 1$, $\beta_{2,1}$ becomes the univariate kurtosis β_2. When $p = 2$, for a two-dimensional random vector $W = (X, Y)^T$, the two-dimensional kurtosis is given by

$$\beta_{2,2} = \frac{\gamma_{4,0} + \gamma_{0,4} + 2\gamma_{2,2} + 4\rho(\rho\gamma_{2,2} - \gamma_{1,3} - \gamma_{3,1})}{(1 - \rho^2)^2} \tag{2}$$

where $\gamma_{k,l}$, $(k, l = 0, 1, 2, 3, 4)$ is

$$\gamma_{k,l} = \frac{E[(X - \mu_x)^k (Y - \mu_y)^l]}{\sigma_x^k \sigma_y^l} \tag{3}$$

and μ_x and μ_y are the marginal means and σ_x and σ_y are the marginal standard deviations and ρ is the correlation between X and Y. In particular, when a two-dimensional random vector W has a discrete probability distribution $f(x_j, y_k)$, $j = 0, 1, ..., n$ and $k = 0, 1, .., m$, the two dimensional kurtosis can be calculated similarly with

$$\gamma_{k,l} = \frac{\sum_{j=0}^{n} \sum_{k=0}^{m} f(x_j, y_k)(x_j - \mu_x)^k (y_k - \mu_y)^l}{\sigma_x^k \sigma_y^l}. \tag{4}$$

From (3) or (4), the marginal kurtoses of the marginal distribution of X and Y are

$$\beta_{2,x} = \gamma_{4,0} \text{ and } \beta_{2,y} = \gamma_{0,4} \tag{5}$$

From Priestley (1981), the normalized spectral density with zeros frequency at the center can be treated as a probability density function. The kurtosis of the spectral density estimate corresponding to (1) of an SEM image can be calculated by (2) and (3) or (4). A sharper SEM image corresponds to a spectrum which has a large shoulder or has a flatter shape. Thus, it can be concluded that the corresponding kurtosis of the sharper image is smaller. Therefore, an increase in kurtosis over some preestablished reference portends that the sharpness of an SEM image has been degraded relative to the existing at the time the reference value was established.

The marginal kurtoses defined in (5) are used to measure the shapes of the marginal distributions. The difference between the marginal kurtoses is a measure of the shape difference between the marginal distributions. It can be used to detect possible instrument vibration. To show how the kurtosis can be applied to the SEM images, a series of five micrographs were taken on the same instrument with only one parameter changed for each image. Two-dimensional kurtoses were calculated for the five micrographs. The result shows the kurtosis measure gives the correct ranking of the sharpness of the images.

Keywords: Image analysis, kurtosis, discrete Fourier transform, spatial frequency.

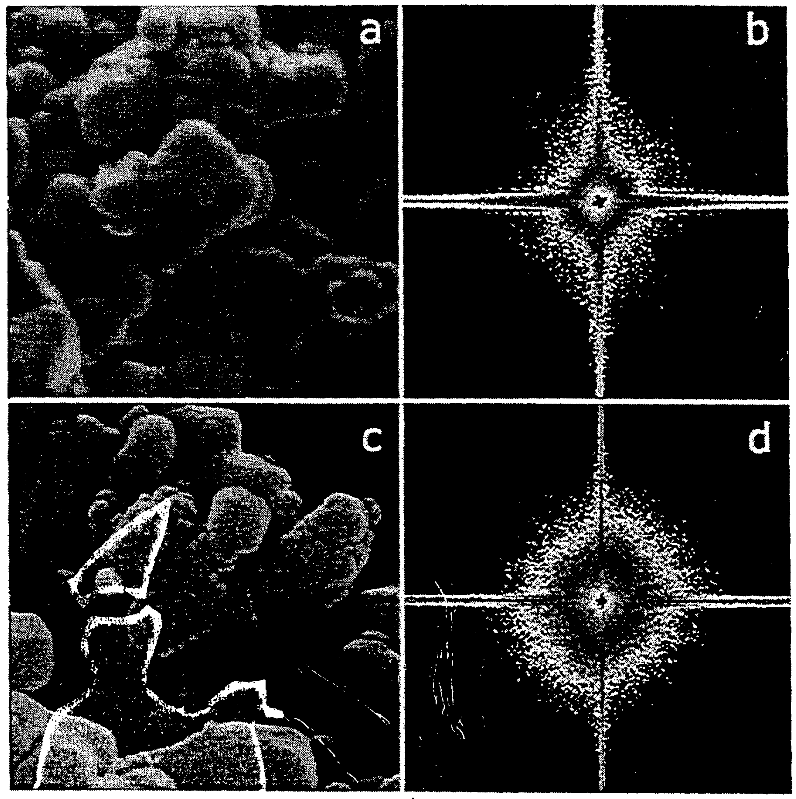

FIGURE 1. Scanning electron micrographs and their two-dimensional Fourier frequency magnitude distributions. (a) Scanning electron micrograph of heavily coated oxide taken following a tip change. (b) Two-dimensional Fourier frequency magnitude distributions of image (a). (c) Scanning electron micrograph of the same sample as above, but taken when the instrument is functioning at a high level of performance. (d) Two-dimensional Fourier frequency magnitude distributions of image (c). Note that there are more high frequency elements present in the Fourier frequency magnitude distribution from (c)